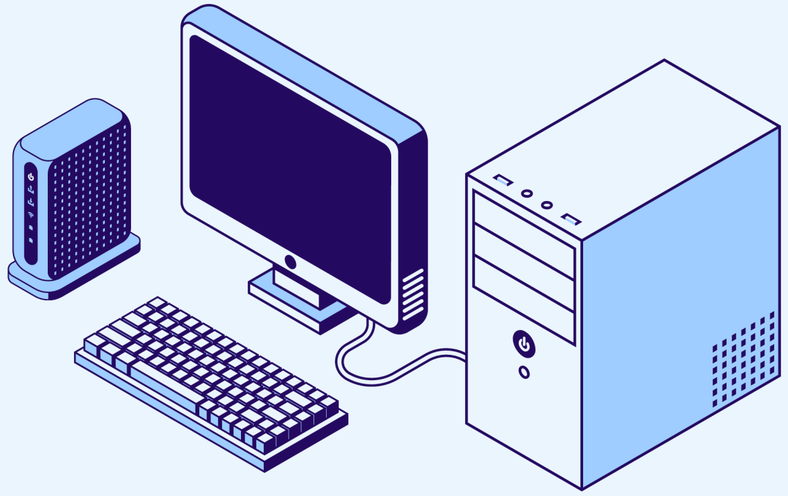
PREPARE FOR DIGITAL PRESERVATION: DIGITAL STORAGE & MANAGEMENT 101

Digital Preservation Training Webinar #3

April 28, 2026 | 12:00PM (PT)



CALIFORNIA REVEALED

GROUNDWORK GRANTS

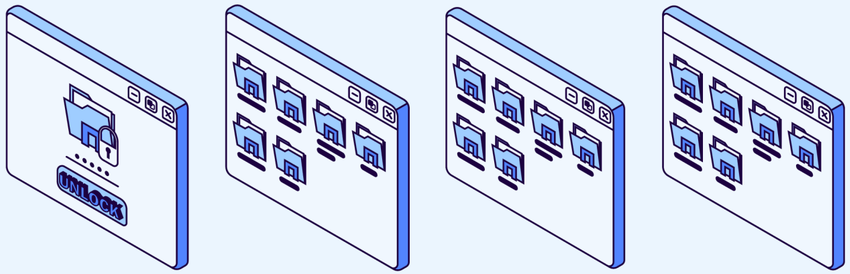
California Cultural Heritage Disaster Preparedness Project

THANK YOU RECOLLECTION WISCONSIN

This series is based on curriculum developed by Recollection Wisconsin, who partnered with UW-Madison and UW-Milwaukee to provide library school students with practical experience in digital stewardship and increase the capacity of small libraries and cultural heritage organizations to curate their digital collections.

Curating Community Digital Collections (CCDC) program and curriculum materials are licensed under CC BY-NC 4.0.
Recollection Wisconsin welcomes the use of these materials to develop your own digital preservation program.

SERIES AGENDA



1 **Tuesday March 31**
Digital Preservation Models & Policy

2 **Tuesday April 14**
Prepare for Digital Preservation: Identifying, Assessing & Prioritizing

3 **TODAY**
Digital Storage & Management 101

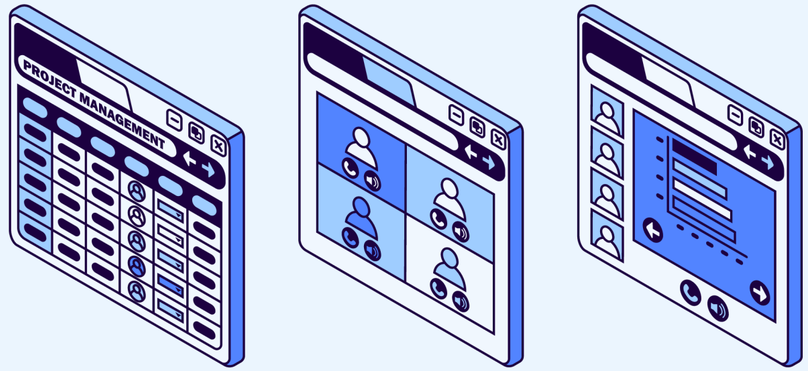
4 **Tuesday May 5**
Digital Collection Development, Outreach & Programming

WEBINAR #3 AGENDA

Here are the topics we will cover this afternoon:

12:00 PM to 1:00 PM: Presentation
1:00 PM to 1:30 PM: Questions, comments, show & tell...

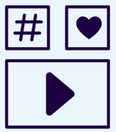
- Metadata, file naming and organizing
- Best(ish) storage practices: 3-2-1
- File checking (aka fixity/integrity)
- Next steps:
  - Play with Advanced Namer
  - Play with CA-R's md5tool
  - Document your Digital Storage 3-2-1 Plan



WHAT CAN WE DO?

1. Digital Preservation policy
- Document your decisions
- Plan for the future

2. Collections Inventory
- Identify what you have
- Prioritize what you need to keep



3. File/folder management scheme
- Get organized! Get the (digital) boxes off the (digital) floor

4. Research and make recommendations for 3-2-1 storage solutions
- Store more than one copy of your files
- Integrity check your data

5. Connect with communities of practice for ongoing support
- You don’t have to go it alone!

Key Definitions:
Digital readiness

Having the knowledge, tools, resources and infrastructure to provide online public access to archives and historical records, and the ability to preserve obsolete audiovisual formats and provide long-term storage of digital records.

Recollection Wisconsin, Digital Toolkit (2023)

Digital Readiness Toolkit (2023) Recollection Wisconsin

Focus Areas and Levels

- **Planning and prioritization**: What is your project plan and how will you decide what to digitize?
- **Permissions and copyright**: How will you know if you can share these materials?
- **Digitization**: How will you undertake the actual digitization work?
- **Description**: How will you describe the digitized materials?
- **Sharing and access**: How will you share what you've digitized?
- **Storage and maintenance**: What is the long-term plan for storing digital materials?
- **Evaluation**: How will you know that what you set out to do was achieved?

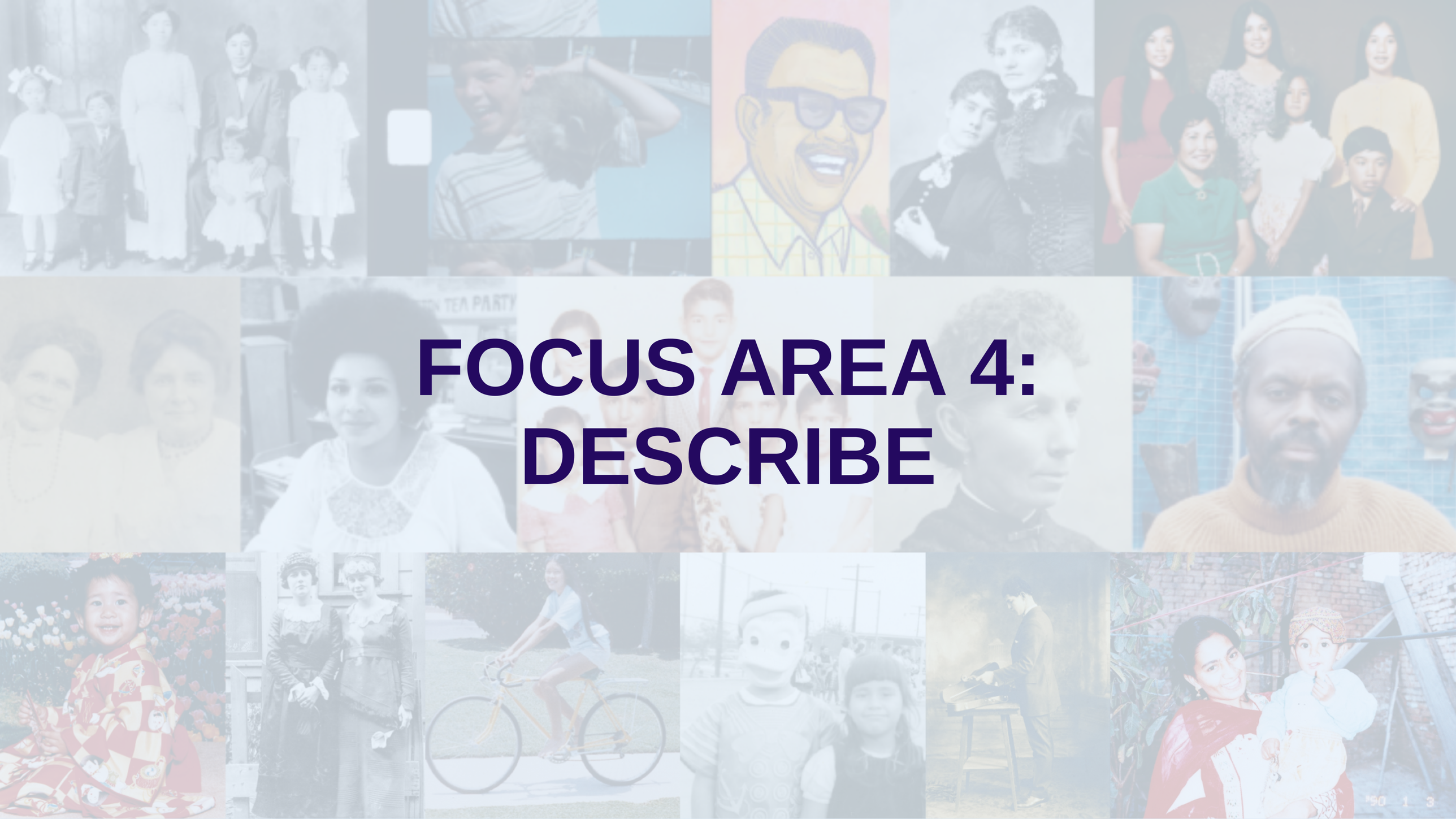
DISASTER RECOVERY

How does digital preservation fit into your overall disaster plan or emergency preparedness plan?

Define the steps that should happen during and in response to natural and human-made disasters that affect digital storage:

- What are all the types of natural disasters, human error, or media failure that could occur?
- What precautions do you take to make sure digital files are protected and prepared for disasters?
- Do you have a system of notification in place? If so, where is it located? If not, who will be included in the notification tree and who will be responsible for its maintenance?
- How will you check if data is lost after a disaster event? If data is lost, how will you recover from backups? What is the procedure for verifying fixity after recovery?

Disaster Preparedness for Neighbors: Introducing Citizen Emergency Response Teams (1991)
City of Huntington Beach Archives



FOCUS AREA 4: DESCRIBE

Key Definitions: Metadata

“Data about data.” Information which describes significant aspects of a resource. Metadata enables successful management and preservation of digital materials over time. It will assist in ensuring essential contextual, historical, and technical information are preserved along with the digital object.

Digital Preservation Handbook Glossary (2026)

TYPES OF METADATA

Administrative

Descriptive

Technical

Preservation

Administrative

Information needed to responsibly manage the object or collection. Includes provenance, rights information and security issues.

unique identifier - old identifier - box numbers - date of production - location - notes

Dolan Sacto Cable #2 (1982)
Center for Sacramento History

Descriptive

About the content. Intended to facilitate finding and identifying the resource.

title - collection - series - creator - date of production - description

Dolan Sacto Cable #2 (1982)
Center for Sacramento History

Technical

Information on how the item was created and/or stored.

physical objects:
media type/format -
duration/dimensions - generation

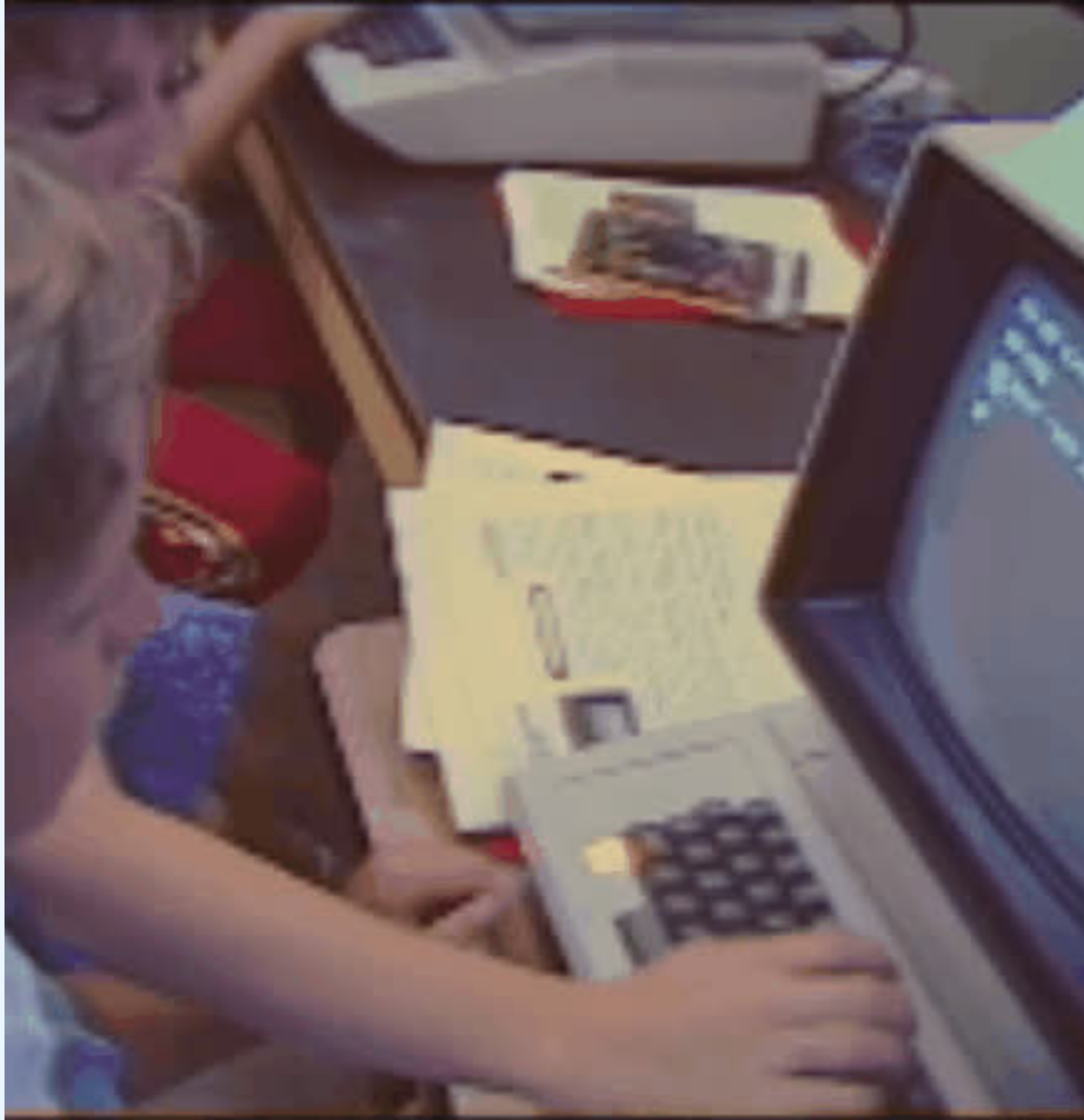
digital objects:
file name - date created - codec -
wrapper -size (bytes)

Dolan Sacto Cable #2 (1982)
Center for Sacramento History

Preservation

Keep track of what's required to maintain and preserve items, and what has already been done.

item condition - container condition - digital copies - preservation actions - preservation technical environment



Dolan Sacto Cable #2 (1982)
Center for Sacramento History

RECOMMENDED FIELDS (ITEM-LEVEL)

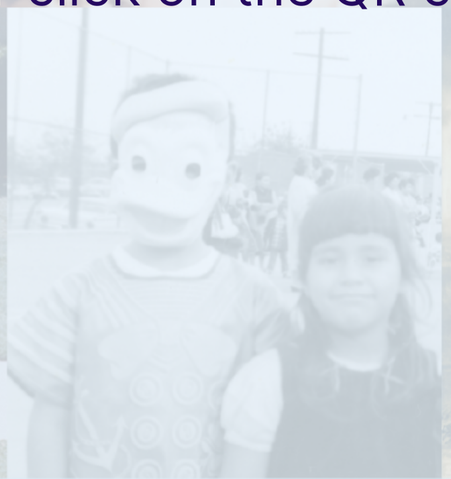
- Organization
- Title
- Object ID
- Created Date
- Creator(s)
- Significance/Description
- Permissions
- Condition
- Material Type
- Format
- Extent/Number of Parts
- Dimensions/Duration
- Location

RECOMMENDED FIELDS (COLLECTION-LEVEL)

- Organization
- Collection Title(s)
- Range of Object IDs
- Created Date(s)
- Creator(s)
- Significance/Description
- Permissions
- Condition
- Material Type
- Format
- Number of Parts (Folders/Files)
- Size of collection (Linear Feet/GB/TB)
- Location (Room/Shelf/Server/Drive)

WHAT RESOURCES CAN YOU LEVERAGE FOR METADATA CREATION?

Consider existing records for both physical and/or digital collections in your care.
What kinds of description can you build upon?
Follow the link in the chat or click on the QR code to respond...



File Naming Tips

Be unique

Be consistent

For your unique identifier (UID) elect a naming scheme & use it across all digitized collections

3-letter institution name abbreviation + 3-letter collection name abbreviation + serialized numbering up to 9,999 + extension

Alexander Media Services Collection = AMSC
Serialized number = 0001
CSH_AMC_0001.tif

Format dates

When using dates in a file name, always start yyyy, then mm, then dd. This ensures that files sort appropriately.

2026-04-14.tif
2026-04-15.tif
2026-05-05.tif

File Naming Tips

Serialization & zero padding

Pad serialization of file names with enough zeros for your collection to ensure files sort correctly.

NOT THIS:	THIS:
1.tif	001.tif
11.tif	002.tif
12.tif	011.tif
2.tif	012.tif

Masters vs. derivatives

It’s good to indicate whether a file is a master or a derivative, although it could be obvious from the file format.

CSH_AMC_0001.tif	casacsh_002331_prsv.mov
CSH_AMC_0001.jpg	casacsh_002331_access.mp4
CSH_AMC_0001.xml	casacsh_002331_PBCore.xml

File Naming Don’ts

Don’t develop a file name dependency

File names are not actually part of the file, but rather part of the file system and are therefore not dependable as being persistent over time and across systems.

The Unique ID (UID) assigned to the object should be the constant identifier used to track and maintain the provenance of the file.

The UID may be the same as the file name, but whatever the case be sure to associate and document the UID appropriately.

Don’t overthink

Think: “Is this logical? Can I spell out the rules easily enough to do batch renaming?”

Whether the filename is a randomly generated value or not, be systematic. In trying to create the perfectly contained and expressed filename or UID structure, there is a strong temptation to overthink them to the point that they become non-systematic or too idiosyncratic to be logically parsed.

File Naming Don’ts

Don’t use filenames as catalog records

Filenames are not the place to cram in a bunch of descriptive and structural information. That’s what databases and spreadsheets are for! Trying to cram excessive descriptive information into a filename:

- creates unwieldy names
- is often futile because of how often conditions or conventions change and new scenarios come up over time.

Don’t make it machine un-readable

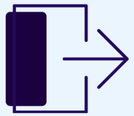
There is often an urge to make a file naming structure decodable by humans, but it also needs to be decodable by computers.

Avoid characters that:

- are not URL compatible (e.g., spaces)
- require escape characters (e.g., parentheses, quotation marks)
- are reserved by operating systems (e.g., colons, slashes)

Limit options to numbers, letters, hyphens, and underscores (and periods before the format suffix)

REMEMBER TO DOCUMENT!



Write down your organization’s file naming rules and share them widely.
Ask all personnel to follow the guidelines.

FOCUS AREA 6: STORE & MAINTAIN

Key Definitions:
Digital Preservation

Refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. Digital preservation refers to all of the actions required to maintain access to digital materials beyond the limits of media failure or technological and organizational change.

Digital Preservation Handbook Glossary (2026)

Digital Storage 101 Overview

1. Without storage, there is no digital preservation!
2. Storage should be planned and budgeted for in cycles
3. Storage/backup goal: 3 copies, 2 types of media, 1 offsite
4. Preservation storage should be reliable and auditable
5. Preservation storage entails requirements beyond backups
6. Remember bestish practices and do your best given your resources!

Baseline Truths

Digital preservation is a relay race, not a marathon.

You’re not looking for the Holy Grail storage medium that fits every situation, never fails, and lasts forever. Not least because it does not exist and never will.

You’re looking to “refresh” your storage every so often. And you need to budget equipment money and staff time accordingly.

Silver lining: Storage tends to lessen in cost over time.
(For now, anyway.)

Sparks - Winner of Race (1921)
Mill Valley Public Library
Lucretia Little History Room

Baseline Truths

The everlasting digital storage medium?

It’s been tried. Every time, it didn’t capture enough market share to survive economically.

Eventually its equipment stopped existing. So much for the Holy Grail.

Anybody promising this is selling snake oil. Do not buy!

Winner of Dipsea Race (1921)
Mill Valley Public Library
Lucretia Little History Room.

3-2-1 PRINCIPLE

LOCKSS

3

Three copies

2

On at least two different storage media

Examples: a desktop/laptop, an external hard drive, a server, a tape backup system, cloud storage

1

With one copy offsite

Ideally in a geographically separate location so a local disaster (fire, tornado, flood) doesn’t wipe out every copy of your data.

BACKUPS

Reliability and **auditability** requirements are central.

- A backup system that doesn’t work might as well not exist. Yearly integrity checks are the absolute minimum acceptable.
- A backup system so complicated or cumbersome that nobody uses it is also useless.
- Speed of access? Unimportant.

Try to use **a different type of storage medium for backups than for working data**. Media diversity reduces risk of data loss:

- HDD & SSD (portable drive - external drive - DAS & NAS - flash memory)
- Magnetic tape (Linear Tape Open/LTO)
- Cloud (Amazon Web Services; Google Cloud; Azure)
- Hosting sites (Internet Archive; YouTube; California Revealed)
- What’s next?

CLOUD RISKS

Business risks
What if they go out of business, or get sued and have their servers confiscated?

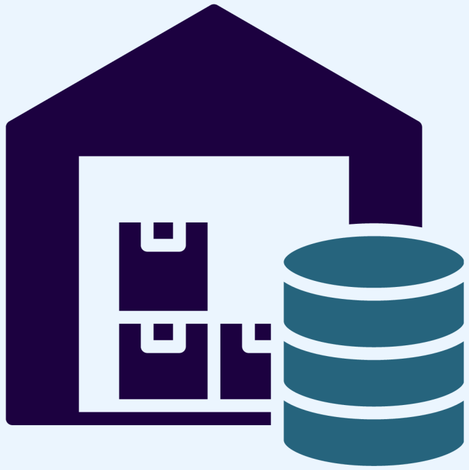
Security risks
Theirs AND Yours
(Yours: Not being careful enough with access permissions such that sensitive data leaks)

Cost risks
Once you get above free tiers…
…you pay per GB/TB stored…
…but you ALSO pay for transfer bandwidth on the network.

This cost can add up!

This doesn’t mean not to use cloud storage. Use it!

It means don’t use it as your only backup. 3-2-1!



HOW MUCH STORAGE?

Part of the answer is **time horizon**.

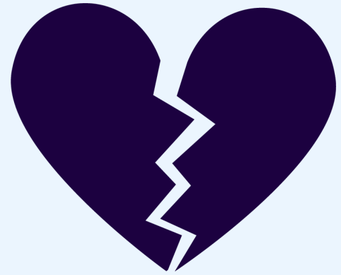
Since digital curation is a relay race, you'll be rethinking the storage issue in a few years anyway. So don't think beyond that rethinking - you'll overspend.

Back-of-the-envelope calculation

- What you have + (how many items you believe you will have x amount of space per item) x 3
- The x3 accounts for 3-2-1 storage.
- Then add some extra.
- If you have IT, ask if there are ways to reduce this burden. There might be! But wear your skeptic hat. ("Get rid of all those giant TIFFs; all you need is JPEG, right?")

Goal: Digital files should remain unchanged over time.

Challenge: Digital files can degrade without warning. You can’t tell if a file has changed just by looking at it.



Key Definitions:
Fixity

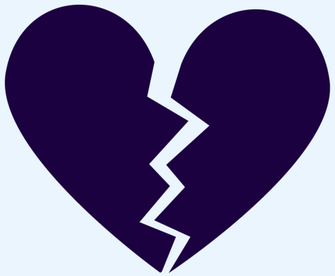
The stability (or fixed-ness) of a digital file.

The knowledge that a digital file is exactly the same as when it was received or created.

Amy Rudersdorf (2020)

Fixity answers the question: *Is the data corrupted or altered?*

Once you have baseline fixity information, you can compare that information with future fixity check information to see if a file has changed or been corrupted.



Fixity answers the question: *Are users getting what they expected?*

Fixity can provide assurance that a copy of a file is unchanged and that what you are providing to your user is exactly what you assert it should be.

Fixity answers the question: *Is something wrong with our system?*

If checks against fixity information for a set of files begin to fail at high rates, it can be an indication of media failure, human error, or other system problems.

SO HOW DO I CHECK MY FILES?

Verify expected file count & total file size

2d4b63341ba16c3ac2cf3008f3865530

Check checksums

WHAT’S A CHECKSUM?

A unique numerical signature dreived from a file. Used to compare copies.

2d4b63341ba16c3ac2cf3008f3865530

Digital Preservation Handbook Glossary (2026)

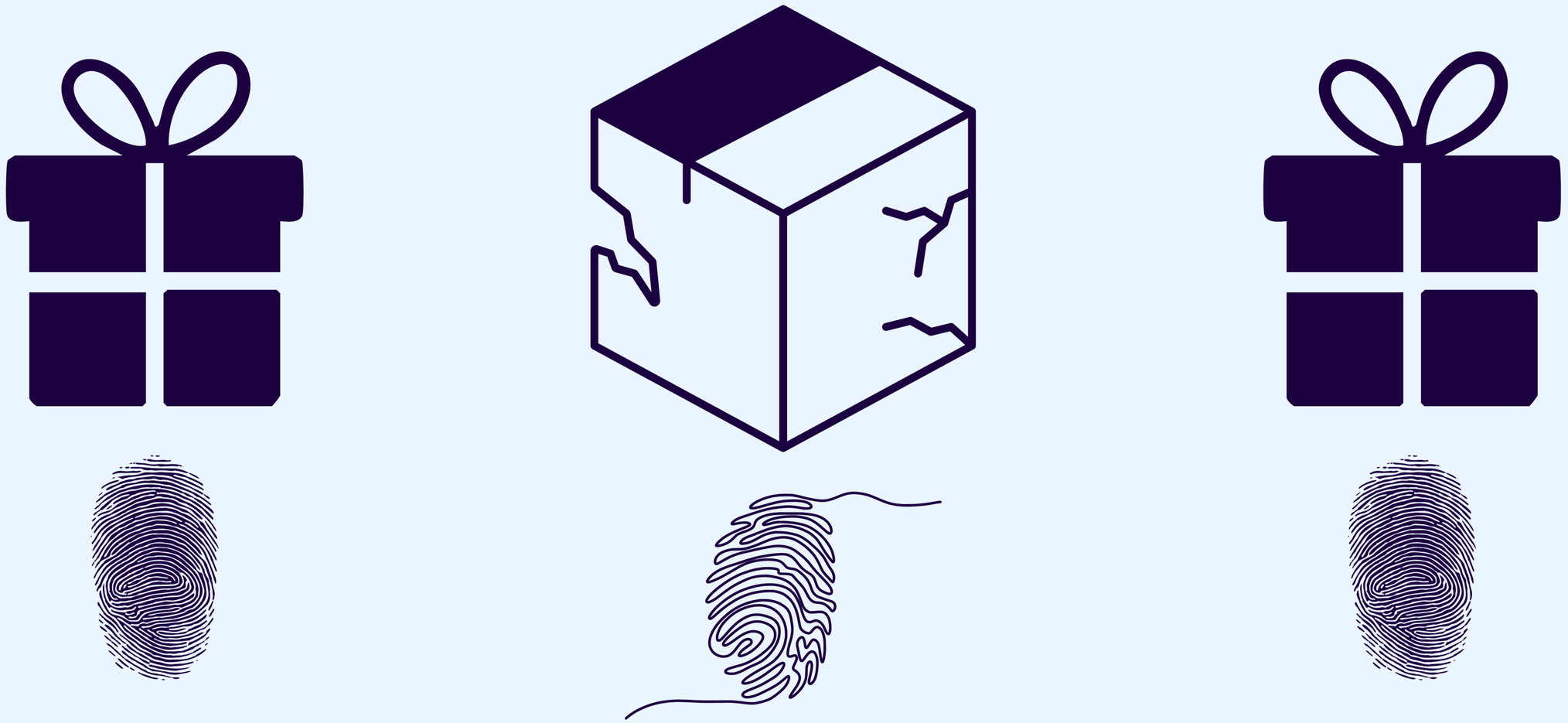
WHY DO WE USE CHECKSUMS?

To ensure the **integrity** of a file after it has been transmitted from one storage device to another.

To confirm a file's **fixity** (it has not degraded or corrupted) after being stored on a device for a period of time.

With some limitations, it can also provide assistance in determining if a file has been modified since it was ingested.

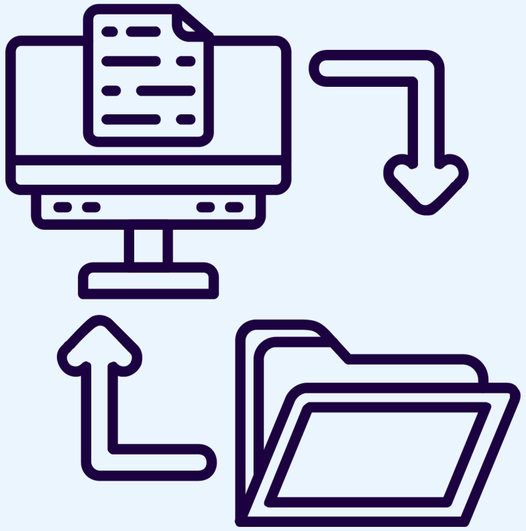
FILE CHECKING (AKA FIXITY/INTEGRITY)



Keep mutliple copies & perform fixity/integrity checks

WHEN SHOULD I CHECK FILES?

When files
are first created
or acquired



When files are
transferred from
one storage system
to another.

At regular
annual
intervals

MD5TOOL DEMO

STORE & MAINTAIN:
KEY ACTIVITIES (BRONZE)

- *Create a collection-level log to document existing and incoming digital collections.*
- *Review and assess current storage methods and locations, including servers, hard drives, thumb drives, etc.*
- *Confirm digital storage estimates for materials you have digitized and plan to digitize.*
- *Review existing digital collections and determine what digital projects are needed:*
  - *Move files onto long term storage or off of unstable hard drives*
  - *Create reference files for playback*
  - *Rename files according to a standard naming convention*
  - *Develop a long term digital preservation strategy*
  - *Determine what digital materials should be kept*
  - *Determine how you will check the integrity of your digital files (fixity)*
  - *Transcode proprietary file formats to an open source file format*

Winner of Girls Race with trophies (1921)
Mill Valley Public Library
Lucretia Little History Room

STORE & MAINTAIN: KEY ACTIVITIES (SILVER)

- *Determine at least one location within your organization where you will store a copy of each unmodified primary file and its related metadata.*
- *Determine at least one location off-site where you will store a copy of each unmodified primary file and its related metadata.*
- *Move copies of files to their storage locations.*
- *Develop a best practice long term storage plan for digital assets.*
- *Develop a best practice back up plan for digital assets*
- *Check and refresh storage media on a regular schedule.*

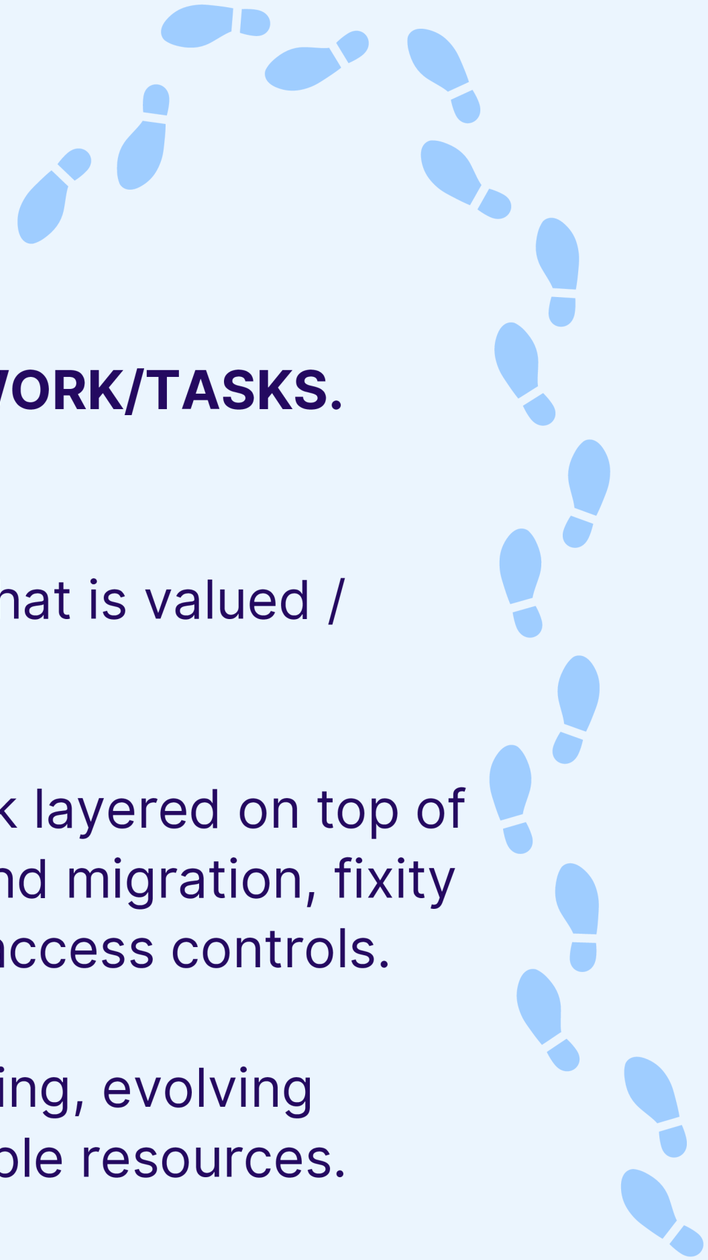
Winner of Girls Race (1921)
Mill Valley Public Library
Lucretia Little History Room

STORE & MAINTAIN: KEY ACTIVITIES (GOLD)

- *Document your storage decisions. Where is it? Who can access it? How?*
- *Determine how you will check the integrity of your digital files (fixity).*
- *Document procedures used for any file checking tools and perform checks on a regular schedule.*
- *Create a schedule and workflow for checking storage locations at least annually.*

Girls Race - Runners (1921)
Mill Valley Public Library
Lucretia Little History Room

FINAL THOUGHTS & NEXT STEPS

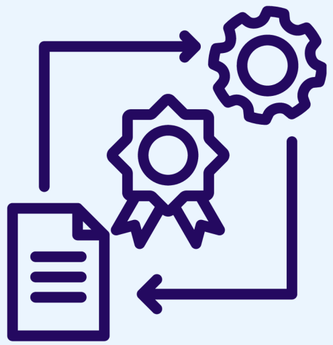


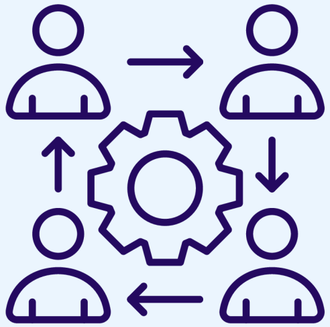
- Relationship between…**POLICY + DECISION-MAKING + WORK/TASKS.** Remember your human resources.

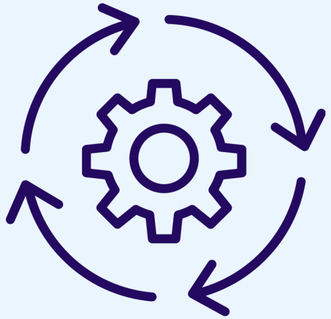
- Keep the process collaborative. Ask your stakeholders: what is valued / mandated / prioritized?

- It’s not the storage medium that counts. It’s the extra work layered on top of the storage: geographic replication, format assessment and migration, fixity checking; logging migrations and deletions; security and access controls.

- Remember the principle of digital stewardship as an ongoing, evolving lifecycle. Bestish practices = doing your best given available resources. Start simple!







HOW IS YOUR DIGITAL PRESERVATION POLICY DEVELOPING?

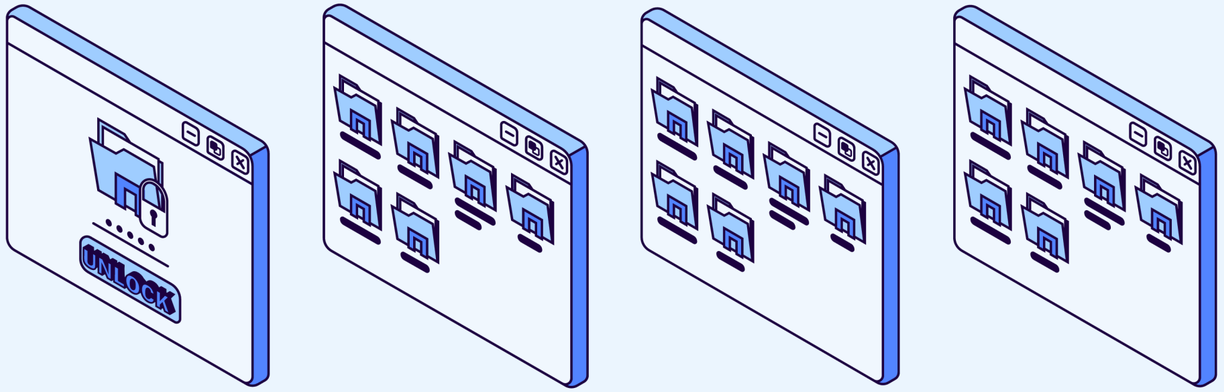
HAVE YOU STARTED THE DIGITAL READINESS SELF-ASSESSMENT?

QUESTIONS? COMMENTS?

Next Webinar: Tuesday, May 5 12PM (PT)
Digital Collection Development, Outreach & Programming

Reminder for California Revealed Partners!
Request your legacy files by May 31, 2026.
californiarevealed.org/digitalpreservationplanning

Join our Digital Preservation Community
groups.google.com/u/2/g/cadigipres

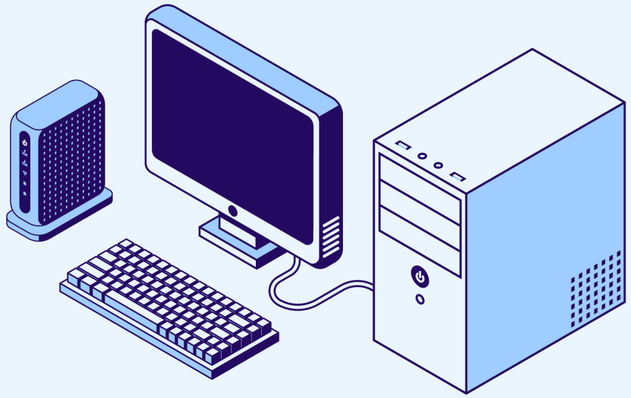


KEEP IN TOUCH!

National City Library (1973)
National City Public Library

team@californiarevealed.org

RESOURCES



Digital Preservation Training Shared Folder: Includes webinar slides and resources

---

Recollection Wisconsin - Digital Readiness Toolkit: Equips public history practitioners — particularly those doing public history work in small, rural, or underrepresented communities, with few or no paid staff — with the foundational knowledge needed to support digital readiness in their organizations.

Digital Project Planning Worksheet: This worksheet aligns with the seven focus areas of Recollection Wisconsin's Digital Readiness Levels. AV Digital Readiness Self-Assessment: the starting point for an organization assessing its digital readiness, with a focus on audiovisual collections.

Digital Preservation Handbook: Brought to you by the Digital Preservation Coalition, this Handbook aims to identify good practice in creating, managing and preserving digital materials and also to provide a range of practical tools to help with that process.

California Revealed Inventory Spreadsheet: To individually list, describe, and assess objects in your collections.

Sustainable Heritage Network (SHN): Managed by the Center for Digital Scholarship and Curation Resources, SHN is a collaborative project that complements the work of Indigenous peoples globally to preserve, share and manage cultural heritage and knowledge.